

c246 Problem Set 3

Due 11 Feb 02, 5pm by email to homework@c246.lbl.gov

1. 15 points

Compare and contrast construction and features of Blosum and PAM series of matrices. Mention strengths and weaknesses of each.

2. 10 points

1-PAM matrix changes an average of 1% of amino acids. Does 2-PAM change on average 2%. Explain.

3. 15 points

Why do we generally use affine gap penalties? What are their drawbacks?

4. 10 points

The best substitution matrix for Smith-Waterman comparisons of distant homologs is often BLOSUM45. The best matrices for BLAST are different; are they higher or lower. Why?

5. 20 points

When will an optimal alignment not be found by FASTA? By Blast?

6. 10 points

Why is it necessary to use masking for low complexity regions?

7. 10 points

It is often valuable to peruse the results of a BLAST search and remove the region of protein that received the most hits, and then search the database again using the remaining regions. Why?

8. 10 points

If looking for common proteins in two as-yet unannotated genomes, why is it better to use tblastx rather than blastn?